CHEMOMETRIC TOOLS FOR AQUAPHOTOMICS

Federico Marini
Department of Chemistry, University of Rome La Sapienza, Rome, Italy
federico.marini@uniroma1.it

Aquaphotomics relies on designing suitable experiments in which the interaction between (infrared) light and aqueous systems constitute the basis for understanding changes in water molecular system, extracting information about the structure of water and identifying possible different water conformations through the recognition of patterns in the recorded absorbance spectra [1]. Therefore, the analysis of highly multivariate spectral data, and the possibility of extracting from them meaningful information by filtering out as much as possible of the unwanted sources of signal variability is a core issue of the data processing pipeline in aquaphotomics. In this context, it is evident how chemometrics (the chemical discipline which makes use of mathematics, statistics, logics and computer science to extract meaningful information from chemical data) may represent an invaluable and needed toolbox [2].

Chemometrics, indeed, accompanies each step of the analytical pipeline for the careful selection of the optimal experimenta conditions under which to carry out the experiments (through a rational experimental design) to the overall processing of the results obtained. Based on these considerations, in the present communication, the main chemometric tools which may be used in the context of aquaphotomics will be illustrated and discussed, by means of selected representative examples.

The importance of an appropriate preprocessing of the data, in order to remove spurious and unwanted sources of variation, which could mask the information of interest will also be stressed and some specific tools, such as the Extended multiplicative scatter correction (EMSC)[3] will be illustrated. Then, the fundamental methods for the exploratory analysis of multivariate data (i.e., principal component analysis (PCA) [4] and multivariate curve resolution (MCR) [5]) will be presented. Lastly, the possibility of implementing predictive modeling, whenever one or more responses are available, by means of multivariate regression techniques, such as principal component regression (PCR) [6] or partial least squares regression (PLS) [6] will also be discussed.

[1] R. Tsenkova, Z. Kovacs, Y. Kubota, Subcell. Biochem. 71 (2015) 189-211.
[2] R. Tsenkova, J. Muncan, B. Pollner, Z. Kovacs, Front. Chem. 6 (2018) 363.
[3] H. Martens, E. Stark, J. Pharm. Biomed. Anal. 9 (1991) 625-635.
[4] R. Bro, A.K. Smilde, Anal. Meth. 6 (2014) 2812-2831.
[5] A. De Juan, J. Jaumot, R. Tauler, Anal. Meth. 6 (2014) 4964-4976.
[6] H. Martens, T. Naes, Multivariate Calibration, 2nd ed., John Wiley and Sons, New York, NY, 1991.